

M-1-E

(2)

0400
04-18-01

#3

OIEP

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/769,066

DATE: 04/26/2001

TIME: 17:06:08

Input Set : N:\Crf3\RULE60\09769066.txt

Output Set: N:\CRF3\04262001\I769066.raw

ENTERED

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Fuerst, Thomas R.
6 McAtee, C. Patrick
7 Yarbough, Patrice O.
8 Zhang, Yifan

10 (ii) TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR

12 (iii) NUMBER OF SEQUENCES: 31

14 (iv) CORRESPONDENCE ADDRESS:

15 (A) ADDRESSEE: Dehlinger & Associates
16 (B) STREET: 350 Cambridge Ave., Suite 250
17 (C) CITY: Palo Alto
18 (D) STATE: CA
19 (E) COUNTRY: USA
20 (F) ZIP: 94306

22 (v) COMPUTER READABLE FORM:

23 (A) MEDIUM TYPE: Floppy disk
24 (B) COMPUTER: IBM PC compatible
25 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
26 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

28 (vi) CURRENT APPLICATION DATA:

C--> 29 (A) APPLICATION NUMBER: US/09/769,066

C--> 30 (B) FILING DATE: 24-Jan-2001

31 (C) CLASSIFICATION:

33 (vii) PRIOR APPLICATION DATA:

34 (A) APPLICATION NUMBER: 08/542,634
35 (B) FILING DATE:

38 (viii) ATTORNEY/AGENT INFORMATION:

39 (A) NAME: Fabian, Gary R.
40 (B) REGISTRATION NUMBER: 33,875
41 (C) REFERENCE/DOCKET NUMBER: 4600-0293.30

43 (ix) TELECOMMUNICATION INFORMATION:

44 (A) TELEPHONE: (415) 324-0880
45 (B) TELEFAX: (415) 324-0960

47 (2) INFORMATION FOR SEQ ID NO: 1:

49 (i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 2049 base pairs
51 (B) TYPE: nucleic acid

W--> 60 (C) STRANDEDNESS: Hepatitis E Virus (Burma strain)

61 ORF-2

53 (D) TOPOLOGY: linear

55 (ii) MOLECULE TYPE: DNA (genomic)

57 (iii) HYPOTHETICAL: NO

59 (vi) ORIGINAL SOURCE:

64 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

66 ATGCGCCCTC GGCCTATTTT GTTGCTGCTC CTCATGTTTT TGCCTATGCT GCCCGCGCCA 60

68 CCGCCCGGTC AGCCGCTGTC CCGCCGTCGT GGGCGGCGCA GCGGCGGTTC CGGCGGTGGT 120

RAW SEQUENCE LISTING

DATE: 04/26/2001

PATENT APPLICATION: US/09/769,066

TIME: 17:06:08

Input Set : N:\Crif3\RULE60\09769066.txt

Output Set: N:\CRF3\04262001\I769066.raw

```

70 TTCTGGGGTG ACCGGGTTGA TTCTCAGCCC TTCGCAATCC CCTATATTCA TCCAACCAAC 180
72 CCCTTCGCCC CCGATGTCAC CGCTGCGGCC GGGGCTGGAC CTCGTGTTTCG CCAACCCGCC 240
74 CGACCACTCG GCTCCGCTTG GCGTGACCAG GCCCAGCGCC CCGCCGTTGC CTCACGTCGT 300
76 AGACCTACCA CAGCTGGGGC CGCGCCGCTA ACCGCGGTCG CTCCGGCCCA TGACACCCCG 360
78 CCAGTGCCCTG ATGTCGACTC CCGCGGCGCC ATCTTGCGCC GGCAGTATAA CCTATCAACA 420
80 TCTCCCCTTA CCTCTTCCGT GGCCACCGGC ACTAACCTGG TTCTTTATGC CGCCCCCTCTT 480
82 AGTCCGCTTT TACCCCTTCA GGACGGCACC AATACCCATA TAATGGCCAC GGAAGCTTCT 540
84 AATTATGCCC AGTACCGGGT TGCCCGTGCC ACAATCCGTT ACCGCCCGCT GGTCCCAAT 600
86 GCTGTGCGCG GTTACGCCAT CTCCATCTCA TTCTGGCCAC AGACCACCAC CACCCCGACG 660
88 TCCGTTGATA TGAATTCAAT AACCTCGACG GATGTTTCGTA TTTTAGTCCA GCCCGGCATA 720
90 GCCTCTGAGC TTGTGATCCC AAGTGAGCGC CTACACTATC GTAACCAAGG CTGGCGCTCC 780
92 GTCGAGACCT CTGGGGTGGC TGAGGAGGAG GCTACCTCTG GTCTGTATTAT GCTTTGCATA 840
94 CATGGCTCAC TCGTAAATTC CTATACTAAT ACACCCATA CCGGTGCCCT CGGGCTGTTG 900
96 GACTTTGCCC TTGAGCTTGA GTTTCGCAAC CTTACCCCGG GTAACACCAA TACGCGGGTC 960
98 TCCCGTTATT CCAGCACTGC TCGCCACCGC TTTCGTGCGG GTGCGGACGG GACTGCCGAG 1020
100 CTCACCACCA CGGCTGCTAC CCGCTTTATG AAGGACCTCT ATTTTACTAG TACTAATGGT 1080
102 GTCGGTGAGA TCGGCCGCGG GATAGCCCTC ACCCTGTTCA ACCTTGCTGA CACTCTGCTT 1140
104 GCGGCGCTGC CGACAGAAAT GATTTCTGCG GCTGGTGGCC AGCTGTTCTA CTCCCGTCCC 1200
106 GTTGTCTCAG CCAATGGCGA GCCGACTGTT AAGTTGTATA CATCTGTAGA GAATGCTCAG 1260
108 CAGGATAAGG GTATTGCAAT CCCGCATGAC ATTGACCTCG GAGAATCTCG TGTGGTTATT 1320
110 CAGGATTATG ATAACCAACA TGAACAAGAT CCGCCGACGC CTTCTCCAGC CCCATCGCGC 1380
112 CCTTTCTCTG TCCTTCGAGC TAATGATGTG CTTTGGCTCT CTCTCACC GC TGCCGAGTAT 1440
114 GACCAGTCCA CTTATGGCTC TTCGACTGGC CCAGTTTATG TTTCTGACTC TGTGACCTTG 1500
116 GTTAATGTTG CGACCGGCGC GCAGGCCGTT GCCCGTTCGC TCGATTGGAC CAAGGTCACA 1560
118 CTTGACGGTC GCGCCCTCTC CACCATCCAG CAGTACTCGA AGACCTTCTT TGTCTGCGC 1620
120 CTCCGCGGTA AGCTCTCTTT CTGGGAGGCA GGCACAATA AAGCCGGGTA CCCTTATAAT 1680
122 TATAACACCA CTGCTAGCGA CCAACTGCTT GTCGAGAATG CCGCCGGGCA CCGGGTCGCT 1740
124 ATTTCCACTT ACACCACTAG CCTGGGTGCT GGTCCGCTCT CCATTCTGCG GGTGCGCGTT 1800
126 TTAGCCCCC ACTCTGCGCT AGCATTGCTT GAGGATACCT TGGACTACCC TGCCCGCGCC 1860
128 CATACTTTTG ATGATTTCTG CCCAGAGTGC CGCCCTTGG GCCTTCAGGG CTGCGCTTTC 1920
130 CAGTCTACTG TCGCTGAGCT TCAGCGCCTT AAGATGAAGG TGGGTAAAAC TCGGGAGTTG 1980
132 TAGTTTATTT GCTTGTGCC CCCTTCTTTC TGTTGCTTAT TTCTCATTTT TCGGTTCCGC 2040
134 GCTCCCTGA 2049

```

136 (2) INFORMATION FOR SEQ ID NO: 2:

138 (i) SEQUENCE CHARACTERISTICS:

139 (A) LENGTH: 2058 base pairs

140 (B) TYPE: nucleic acid

W--> 149 (C) STRANDEDNESS: Hepatitis E Virus (Mexico Strain)

150 ORF-2 region

142 (D) TOPOLOGY: linear

144 (ii) MOLECULE TYPE: DNA (genomic)

146 (iii) HYPOTHETICAL: NO

148 (vi) ORIGINAL SOURCE:

152 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

154 ATGCGCCCTA GGCTCTTTT GCTGTTGTTT CTCTGTTTTC TGCCTATGTT GCCCGCGCCA 60
156 CCGACCGGTC AGCCGTCTGG CCGCGTCTCG GGGCGGCGCA GCGGCGGTAC CCGCGGTGGT 120
158 TTCTGGGGTG ACCGGGTTGA TTCTCAGCCC TTCGCAATCC CCTATATTCA TCCAACCAAC 180
160 CCCTTTGCCC CAGACGTTGC CGCTGCGTCC GGGTCTGGAC CTCGCCTTCG CCAACCAGCC 240
162 CGGCCACTTG GCTCCACTTG GCGAGATCAG GCCCAGCGCC CCTCCGCTGC CTCCCGTCGC 300

```

RAW SEQUENCE LISTING

DATE: 04/26/2001

PATENT APPLICATION: US/09/769,066

TIME: 17:06:08

Input Set : N:\Crif3\RULE60\09769066.txt

Output Set: N:\CRF3\04262001\I769066.raw

```

164 CGACCTGCCA CAGCCGGGGC TCGGGCGCTG ACGGCTGTGG CGCCTGCCCA TGACACCTCA 360
166 CCCGTCCCGG ACGTTGATTC TCGCGGTGCA ATTCTACGCC GCCAGTATAA TTTGTCTACT 420
168 TCACCCCTGA CATCCTCTGT GGCCTCTGGC ACTAATTTAG TCCTGTATGC AGCCCCCTT 480
170 AATCCGCCTC TGCCGCTGCA GGACGGTACT AATACTCACA TTATGGCCAC AGAGGCCTCC 540
172 AATTATGCAC AGTACCGGGT TGCCCCGCGT ACTATCCGTT ACCGGCCCCCT AGTGCCTAAT 600
174 GCAGTTGGAG GCTATGCTAT ATCCATTTCT TTCTGGCCTC AAACAACCAC AACCCCTACA 660
176 TCTGTTGACA TGAATTCAT TACTTCCACT GATGTCAGGA TTCTTGTTCA ACCTGGCATA 720
178 GCATCTGAAT TGGTCATCCC AAGCGAGCGC CTTCACTACC GCAATCAAGG TTGGCGCTCG 780
180 GTTGAGACAT CTGGTGTGTC TGAGGAGGAA GCCACCTCCG GTCTTGTCAT GTTATGCATA 840
182 CATGGCTCTC CAGTTAACTC CTATACCAAT ACCCCTTATA CCGGTGCCCT TGGCTTACTG 900
184 GACTTTGCCT TAGAGCTTGA GTTTCGCAAT CTCACCACCT GTAACACCAA TACACGTGTG 960
186 TCCCGTTACT CCAGCACGGC CCGTCACCGG CTCCGCCGAG GGGCCGACGG GACTGCGGAG 1020
188 CTGACCACAA CTGCAGCCAC CAGGTTTCATG AAAGATCTCC ACTTTACCGG CCTTAATGGG 1080
190 GTAGGTGAAG TCGGCCGCGG GATAGCTCTA ACATTACTTA ACCTTGCTGA CACGCTCCTC 1140
192 GCGGGGCTCC CGACAGAATT AATTTCTGTC GCTGGCGGGC AACTGTTTTA TCCCGCCCCG 1200
194 GTTGCTCTCAG CCAATGGCGA GCCAACCGTG AAGCTCTATA CATCAGTGGA GAATGCTCAG 1260
196 CAGGATAAGG GTGTTGCTAT CCCCCACGAT ATCGATCTTG GTGATTCGCG TGTGGTCATT 1320
198 CAGGATTATG ACAACCAGCA TGAGCAGGAT CGGCCACCCC CGTCGCCTGC GCCATCTCGG 1380
200 CCTTTTTCTG TTCTCCGAGC AAATGATGTA CTTTGCTGTG CCCTCACTGC AGCCGAGTAT 1440
202 GACCAGTCCA CTTACGGGTC GTCAACTGGC CCGGTTTATA TCTCGGACAG CGTGACTTTG 1500
204 GTGAATGTTG CGACTGGCGC GCAGGCCGTA GCCCGATCGC TTGACTGGTC CAAAGTCACC 1560
206 CTCGACGGGC GGCCCCCTCC GACTGTTGAG CAATATTCCA AGACATTCTT TGTGCTCCCC 1620
208 CTTCGTGGCA AGCTCTCCTT TTGGGAGGCC GGCACAACAA AAGCAGGTTA TCCTTATAAT 1680
210 TATAATACTA CTGCTAGTGA CCAGATTCTG ATTGAAAATG CTGCCGGCCA TCGGGTCGCC 1740
212 ATTTCAACCT ATACCACCAG GCTTGGGGCC GGTCCGGTCG CCATTCTGCG GGCCGCGGTT 1800
214 TTGGCTCCAC GCTCCGCCCT GGCTCTGCTG GAGGATACTT TTGATTATCC GGGGCGGGCG 1860
216 CACACATTTG ATGACTTCTG CCCTGAATGC CGCGCTTTAG GCCTCCAGGG TTGTGCTTTC 1920
218 CAGTCAACTG TCGCTGAGCT CCAGCGCCTT AAAGTTAAGG TGGGTAAAC TCGGGAGTTG 1980
220 TAGTTTATTT GGCTGTGCCC ACCTACTTAT ATCTGCTGAT TTCTTTTATT TCCTTTTCT 2040
222 CGGTCCCGCG CTCCCTGA 2058

```

224 (2) INFORMATION FOR SEQ ID NO: 3:

226 (i) SEQUENCE CHARACTERISTICS:

227 (A) LENGTH: 1647 base pairs

228 (B) TYPE: nucleic acid

229 (C) STRANDEDNESS: double

230 (D) TOPOLOGY: linear

232 (ii) MOLECULE TYPE: DNA (genomic)

234 (iii) HYPOTHETICAL: NO

236 (vi) ORIGINAL SOURCE:

237 (C) INDIVIDUAL ISOLATE: Hepatitis E virus (Burma) r62kDa,

238 FIGURE 2

241 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

243 GCGGTCGCTC CGGCCCATGA CACCCGCCA GTGCCTGATG TCGACTCCCG CGGCGCCATC 60
245 TTGCGCCGGC AGTATAACCT ATCAACATCT CCCCTTACCT CTTCGTTGGC CACCGGCACT 120
247 AACCTGGTTC TTTATGCCGC CCCTCTTAGT CCGCTTTTAC CCCTTCAGGA CGGCACCAAT 180
249 ACCCATATAA TGGCCACGGA AGCTTCTAAT TATGCCAGT ACCGGGTTGC CCGTGCCACA 240
251 ATCCGTTACC GCCCGCTGGT CCCCAATGCT GTCGGCGGTT ACGCCATCTC CATCTCATTC 300
253 TGGCCACAGA CCACCACCAC CCCGACGTCC GTTGATATGA ATTCAATAAC CTCGACGGAT 360
255 GTTCGTATTT TAGTCCAGCC CGGCATAGCC TCTGAGCTTG TGATCCCAAG TGAGCGCCTA 420

```

RAW SEQUENCE LISTING

DATE: 04/26/2001

PATENT APPLICATION: US/09/769,066

TIME: 17:06:08

Input Set : N:\Crf3\RULE60\09769066.txt

Output Set: N:\CRF3\04262001\I769066.raw

```

257 CACTATCGTA ACCAAGGCTG GCGCTCCGTC GAGACCTCTG GGGTGGCTGA GGAGGAGGCT 480
259 ACCTCTGGTC TTGTTATGCT TTGCATACAT GGCTCACTCG TAAATTCCTA TACTAATACA 540
261 CCCTATACCG GTGCCCTCGG GCTGTTGGAC TTTGCCCTTG AGCTTGAGTT TCGCAACCTT 600
263 ACCCCCGGTA ACACCAATAC GCGGGTCTCC CGTTATTCCA GACTGCTCG CCACCGCCTT 660
265 CGTCGCGGTG CGGACGGGAC TGCCGAGCTC ACCACCACGG CTGCTACCCG CTTTATGAAG 720
267 GACCTCTATT TTAAGTAGTAC TAATGGTGTC GGTGAGATCG GCCGCGGGAT AGCCCTCACC 780
269 CTGTTCAACC TTGCTGACAC TCTGCTTGGC GGCCTGCCGA CAGAATTGAT TTCGTGGGCT 840
271 GGTGCCAGC TGTCTACTC CCGTCCGTT GTCTCAGCCA ATGGCGAGCC GACTGTTAAG 900
273 TTGTATACAT CTGTAGAGAA TGCTCAGCAG GATAAGGGTA TTGCAATCCC GCATGACATT 960
275 GACCTCGGAG AATCTCGTGT GGTTATTGAG GATTATGATA ACCAACATGA ACAAGATCGG 1020
277 CCGACGCCTT TTCCAGCCCC ATCGCGCCCT TTCTCTGTCC TTCGAGCTAA TGATGTGCTT 1080
279 TGCTCTCTCT TCACCGCTGC CGAGTATGAC CAGTCCACTT ATGGCTCTTC GACTGGCCCA 1140
281 GTTTATGTTT CTGACTCTGT GACCTTGCTT AATGTTGCGA CCGGCGCGCA GGCCGTTGCC 1200
283 CGGTGCTCG ATTGGACCAA GGTACACTT GACGGTCGCC CCCTCTCCAC CATCCAGCAG 1260
285 TACTCGAAGA CCTTCTTTGT CCGTCCGCTC CGCGGTAAGC TCTCTTTCTG GGAGGCAGGC 1320
287 ACAACTAAAG CCGGGTACCC TTATAATTAT AACACCACTG CTAGCGACCA ACTGCTTGTC 1380
289 GAGAATGCCG CCGGGCACC GGTGCTATT TCCACTTACA CCACTAGCCT GGGTGTGCTT 1440
291 CCCGTCTCCA TTTCTGCGGT TGCCGTTTGA GCGCCCACT CTGCGCTAGC ATTGCTTGAG 1500
293 GATACCTTGG ACTACCCTGC CCGCGCCCAT ACTTTTGATG ATTTCTGCCC AGAGTGCCGC 1560
295 CCCCTTGGCC TTCAGGGCTG CGCTTTCCAG TCTACTGTCT CTGAGCTTCA GCGCCTTAAG 1620
297 ATGAAGGTGG GTAAACTCG GGAGTTG

```

299 (2) INFORMATION FOR SEQ ID NO: 4:

301 (i) SEQUENCE CHARACTERISTICS:

302 (A) LENGTH: 1647 base pairs

303 (B) TYPE: nucleic acid

W--> 312 (C) STRANDEDNESS: Hepatitis E virus (Mexico strain)

313 r62kDa, FIGURE 2

305 (D) TOPOLOGY: linear

307 (ii) MOLECULE TYPE: DNA (genomic)

309 (iii) HYPOTHETICAL: NO

311 (vi) ORIGINAL SOURCE:

316 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```

317 GCTGTGGCGC CTGCCCATGA CACCTCACCC GTCCCGGACG TTGATTCTCG CGGTGCAATT 60
319 CTACGCCGCC AGTATAATTT GTCTACTTCA CCCCTGACAT CCTCTGTGGC CTCTGGCACT 120
321 AATTTAGTCC TGTATGCAGC CCCCCTTAAT CCGCTCTGCG CGCTGCAGGA CGGTACTAAT 180
323 ACTCACATTA TGGCCACAGA GGCTCCAAT TATGCACAGT ACCGGGTTGC CCGCGCTACT 240
325 ATCCGTTACC GGCCCTAGT GCCTAATGCA GTTGGAGGCT ATGCTATATC CATTTCTTTC 300
327 TGGCCTCAAA CAACCACAAC CCCTACATCT GTTGACATGA ATTCCATTAC TTCCACTGAT 360
329 GTCAGGATTC TTGTTCAACC TGGCATAGCA TCTGAATTGG TCATCCCAAG CGAGCGCCTT 420
331 CACTACCGCA ATCAAGGTTG GCGCTCGGTT GAGACATCTG GTGTTGCTGA GGAGGAAGCC 480
333 ACCTCCGGT TTGTCAATG ATGCATACAT GGCTCTCCAG TTAAGTCTTA TACCAATACC 540
335 CCTTATACCG GTGCCCTTGG CTTACTGGAC TTTGCCTTAG AGCTTGAGTT TCGCAATCTC 600
337 ACCACCTGTA ACACCAATAC ACGTGTGTCC CGTTACTCCA GCACGGCCCG TCACCGGCTC 660
339 CGCCGAGGGG CCGACGGGAC TGCGGAGCTG ACCACAATG CAGCCACCAG GTTCATGAAA 720
341 GATCTCCACT TTACCGGCCT TAATGGGGTA GGTGAAGTCG GCCGCGGGAT AGCTCTAACA 780
343 TTAATTAACC TTGCTGACAC GTCCTCGGC GGGCTCCCGA CAGAATTAAT TTCGTGGGCT 840
345 GCGGGGCAAC TGTTTTATTC CCGCCCGGTT GTCTCAGCCA ATGGCGAGCC AACCGTGAAG 900
347 CTCTATACAT CAGTGGAGAA TGCTCAGCAG GATAAGGGTG TTGCTATCCC CCACGATATC 960
349 GATCTTGGTG ATTCGCGTGT GGTCATTGAG GATTATGACA ACCAGCATGA GCAGGATCGG 1020

```

RAW SEQUENCE LISTING

DATE: 04/26/2001

PATENT APPLICATION: US/09/769,066

TIME: 17:06:08

Input Set : N:\Crif3\RULE60\09769066.txt

Output Set : N:\CRF3\04262001\I769066.raw

```

351 CCCACCCCGT CGCCTGCGCC ATCTCGGCCT TTTTCTGTTC TCCGAGCAAA TGATGTACTT 1080
353 TGCTGTGCCC TCACTGCAGC CGAGTATGAC CAGTCCACTT ACGGGTCGTC AACTGGCCCG 1140
355 GTTTATATCT CGGACAGCGT GACTTTGGTG AATGTTGCGA CTGGCGCGCA GGCCGTAGCC 1200
357 CGATCGCTTG ACTGGTCCAA AGTCACCCTC GACGGGCGGC CCCTCCCGAC TGTGAGCAA 1260
359 TATCCAAGA CATTCTTTGT GCTCCCCCTT CGTGGCAAGC TCTCCTTTTGG GGAGGCCGGC 1320
361 ACAACAAAAG CAGGTTATCC TTATAATTAT AATACTACTG CTAGTGACCA GATTCTGATT 1380
363 GAAATGCTG CCGGCCATCG GGTCGCCATT TCAACCTATA CCACCAGGCT TGGGGCCGGT 1440
365 CCGGTCGCCA TTTCTGCGGC CGCGGTTTGT GCTCCACGCT CCGCCCTGGC TCTGCTGGAG 1500
367 GATACTTTTG ATTATCCGGG GCGGGCGCAC ACATTTGATG ACTTCTGCCC TGAATGCCGC 1560
369 GCTTTAGGCC TCCAGGGTTG TGCTTTCCAG TCAACTGTCTG CTGAGCTCCA GCGCCTTAAA 1620
371 GTTAAGGTGG GTAAACTCG GGAGTTG 1647
373 (2) INFORMATION FOR SEQ ID NO: 5:
375     (i) SEQUENCE CHARACTERISTICS:
376         (A) LENGTH: 984 base pairs
377         (B) TYPE: nucleic acid
W--> 386     (C) STRANDEDNESS: Hepatitis E Virus (Burma strain) SG3
387         region
379     (D) TOPOLOGY: linear
381     (ii) MOLECULE TYPE: DNA (genomic)
383     (iii) HYPOTHETICAL: NO
385     (vi) ORIGINAL SOURCE:
390     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
392 GGTGCGGACG GGAAGTCCGA GTCACCACC ACGGCTGCTA CCCGCTTTAT GAAGGACCTC 60
394 TATTTTACTA GTACTAATGG TGTCGGTGAG ATCGGCCGCG GGATAGCCCT CACCCTGTTC 120
396 AACCTTGCTG AACTCTGCTT TGGCGGCTG CCGACAGAAT TGATTTCGTC GGCTGGTGGC 180
398 CAGCTGTTCT ACTCCCGTCC CGTTGTCTCA GCCAATGGCG AGCCGACTGT TAAGTTGTAT 240
400 ACATCTGTAG AGAATGCTCA GCAGGATAAG GGTATTGCAA TCCCGCATGA CATTGACCTC 300
402 GGAGAATCTC GTGTGGTTAT TCAGGATTAT GATAACCAAC ATGAACAAGA TCGGCCGACG 360
404 CCTTCTCCAG CCCCATCGCG CCCTTCTCTT GTCCTTCGAG CTAATGATGT GCTTTGGCTC 420
406 TCTCTCACCG CTGCCGAGTA TGACCAGTCC ACTTATGGCT CTTCGACTGG CCGAGTTTAT 480
408 GTTCTGACT CTGTGACCTT GGTTAATGTT GCGACCGGCG CGCAGGCCGT TGCCCGGTCG 540
410 CTCGATTGGA CCAAGGTCAC ACTTGACGGT CGCCCCCTCT CCACCATCCA GCAGTACTCG 600
412 AAGACCTTCT TTGTCCTGCC GTCGCGCGGT AAGCTCTCTT TCTGGGAGGC AGGCACAAC 660
414 AAAGCCGGGT ACCCTTATAA TTATAACACC ACTGCTAGCG ACCAACTGCT TGTCGAGAAT 720
416 GCCGCCGGGC ACCGGGTCGC TATTTCCACT TACACCACTA GCCTGGGTGC TGGTCCCGTC 780
418 TCCATTTCTG CGGTTGCCGT TTTAGCCCCC CACTCTGCGC TAGCATTGCT TGAGGATACC 840
420 TTGGACTACC CTGCCCCGCG CCATACTTTT GATGATTTCT GCCCAGAGTG CCGCCCCCTT 900
422 GGCCTTCAGG GCTGCGCTTT CCAGTCTACT GTCGCTGAGC TTCAGCGCCT TAAGATGAAG 960
424 GTGGGTAAAA CTCGGGAGTT GTAG 984
426 (2) INFORMATION FOR SEQ ID NO: 6:
428     (i) SEQUENCE CHARACTERISTICS:
429         (A) LENGTH: 984 base pairs
430         (B) TYPE: nucleic acid
W--> 439     (C) STRANDEDNESS: Hepatitis E Virus (Mexico strain) SG3
440         region
432     (D) TOPOLOGY: linear
434     (ii) MOLECULE TYPE: DNA (genomic)
436     (iii) HYPOTHETICAL: NO
438     (vi) ORIGINAL SOURCE:

```

VERIFICATION SUMMARY

DATE: 04/26/2001

PATENT APPLICATION: US/09/769,066

TIME: 17:06:09

Input Set : N:\Crf3\RULE60\09769066.txt

Output Set: N:\CRF3\04262001\I769066.raw

L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:60 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:60 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=1
L:149 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:149 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=2
L:312 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:312 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=4
L:386 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:386 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=5
L:439 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:439 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=6
L:492 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:492 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=7
L:517 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:517 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=8
L:542 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:542 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=9
L:575 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:575 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=10
L:608 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:608 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=11
L:631 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:631 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=12
L:653 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:653 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=13
L:797 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:797 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=14
L:941 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:941 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=15
L:1065 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:1065 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=16
L:1188 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:1188 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=17
L:1269 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:1269 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=18
L:1350 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:1350 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=19
L:1377 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:1377 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=20
L:1404 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:1404 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=21
L:1446 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:1446 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=22
L:1488 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:1488 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=23
L:1516 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:1516 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=24

VERIFICATION SUMMARY

DATE: 04/26/2001

PATENT APPLICATION: US/09/769,066

TIME: 17:06:09

Input Set : N:\Crf3\RULE60\09769066.txt

Output Set: N:\CRF3\04262001\I769066.raw

L:1544 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:1544 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=25
L:1665 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:1665 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=26
L:1785 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:1785 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=27
L:1903 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:1903 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=28
L:2015 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=29
L:2036 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=30
L:2115 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:31
L:2118 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:31